Statistical Issues in Drug Development

Second Edition

Stephen Senn

Department of Statistics
University of Glasgow, UK

John Wiley & Sons, Ltd

This page intentionally left blank

Statistical Issues in Drug Development

Statistics in Practice

Advisory Editor

Stephen Senn
University of Glasgow, UK

Founding Editor

Vic Barnett
Nottingham Trent University, UK

Statistics in Practice is an important international series of texts which provide detailed coverage of statistical concepts, methods and worked case studies in specific fields of investigation and study.

With sound motivation and many worked practical examples, the books show in down-to-earth terms how to select and use an appropriate range of statistical techniques in a particular practical field within each title's special topic area.

The books provide statistical support for professionals and research workers across a range of employment fields and research environments. Subject areas covered include medicine and pharmaceutics; industry, finance and commerce; public services; the earth and environmental sciences, and so on.

The books also provide support to students studying statistical courses applied to the above areas. The demand for graduates to be equipped for the work environment has led to such courses becoming increasingly prevalent at universities and colleges.

It is our aim to present judiciously chosen and well-written workbooks to meet everyday practical needs. Feedback of views from readers will be most valuable to monitor the success of this aim.

A complete list of titles in this series appears at the end of the volume.

Statistical Issues in Drug Development

Second Edition

Stephen Senn

Department of Statistics
University of Glasgow, UK

John Wiley & Sons, Ltd

Copyright © 2007

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester,
West Sussex PO19 8SQ, England

Telephone (+44) 1243 779777

Email (for orders and customer service enquiries): cs-books@wiley.co.uk
Visit our Home Page on www.wiley.com

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except under the terms of the Copyright, Designs and Patents Act 1988 or under the terms of a licence issued by the Copyright Licensing Agency Ltd, 90 Tottenham Court Road, London W1T 4LP, UK, without the permission in writing of the Publisher. Requests to the Publisher should be addressed to the Permissions Department, John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, or emailed to permreq@wiley.co.uk, or faxed to (+44) 1243 770620.

This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the Publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Other Wiley Editorial Offices

John Wiley & Sons Inc., 111 River Street, Hoboken, NJ 07030, USA

Jossey-Bass, 989 Market Street, San Francisco, CA 94103-1741, USA

Wiley-VCH Verlag GmbH, Boschstr. 12, D-69469 Weinheim, Germany

John Wiley & Sons Australia Ltd, 42 McDougall Street, Milton, Queensland 4064, Australia

John Wiley & Sons (Asia) Pte Ltd, 2 Clementi Loop #02-01, Jin Xing Distripark, Singapore 129809

John Wiley & Sons Canada Ltd, 6045 Freemont Blvd, Mississauga, Ontario, L5R 4J3

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Anniversary Logo Design: Richard J. Pacifico

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-470-01877-4

Typeset in 10/12pt Times by Integra Software Services Pvt. Ltd, Pondicherry, India
Printed and bound in Great Britain by Antony Rowe Ltd, Chippenham, Wiltshire
This book is printed on acid-free paper responsibly manufactured from sustainable forestry in which at least two trees are planted for each one used for paper production.

To Victoria, Helen and Mark

This page intentionally left blank

Contents

This page intentionally left blank

Preface to the Second Edition

There have been many developments since the first edition of this book and it was high time for a second. My own period working in the pharmaceutical industry is now a distant memory but the ten years working as an academic since the first edition has had its compensations. I have been fortunate enough to be able to consult for many pharmaceutical companies during this time and this has certainly widened my appreciation of the work that statisticians do within the industry and the problems they face.

Alas, this appreciation is not shared by all. Many take it as almost axiomatic that statistical analysis carried out within the pharmaceutical industry is necessarily inferior to that carried out elsewhere. Indeed, one medical journal has even gone so far as to make it a requirement for publication that analyses from the pharmaceutical industry should be confirmed by an academic statistician, a policy which is as impractical as it is illogical.

Two related developments since the first edition, one of which is personal, are highly relevant. The first is that I have been honoured to succeed Vic Barnett as an editor for Wiley's *Statistics in Practice* series, in which this book appears. The second is that the series itself, so ably founded by Vic and Helen Ramsey, has been growing steadily and since the first edition now has attracted a number of further volumes that are highly relevant to this one. The chapters that have particularly benefited are listed with the relevant references as follows.

Chapter 6	Allocating treatments to patients in clinical trials (Berger, 2005)
Chapter 7	Baselines and covariate information (Berger, 2005)
Chapter 11	Intention to treat, missing data and related matters (Molenberghs and Kenward, 2007)
Chapter 16	Meta-analysis (Whitehead, 2002)
Chapter 19	Sequential trials (Ellenberg, et al., 2003)
Chapter 20	Dose-finding (Chevret, 2006)
Chapter 22	Bioequivalence studies (Hauschke, et al., 2007)
Chapter 23	Safety data, harms, drug-monitoring and pharmacoepidemiology (Lui, 2004)
Chapter 24	Pharmacoeconomics and portfolio management (Parmigiani, 2002, Willan and Briggs, 2006)

Also extremely useful are two books on Bayesian methods (O'Hagan *et al.*, 2006, Spiegelhalter *et al.*, 2003) and Brown and Prescott's book on mixed models (Brown and Prescott, 2006), which is already in a second edition. The books on survival analysis and

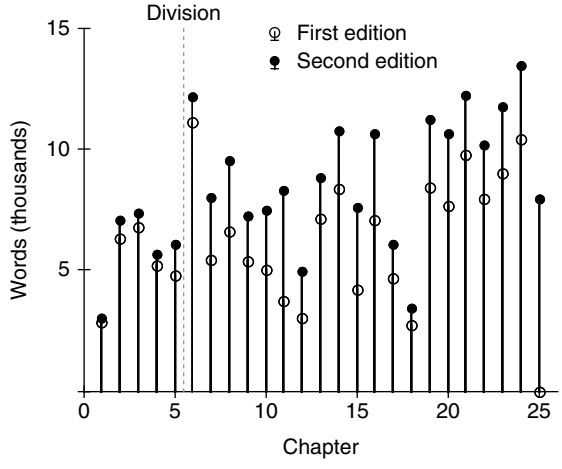

Figure P.1 Comparison of the first and second editions in length.

sequential analysis by Marubini and Valsecchi (1995) and Whitehead (1997) remain highly relevant, of course, as does my own on cross-over trials, (Senn, 2002), also now in a second edition.

All chapters have been brought up to date in the new edition and in particular, there is extensive reference to various guidelines of the International Conference of Harmonization that have been issued since the first edition, in particular ICH E9, *Statistical Principles for Clinical Trials*. A new chapter on pharmaco-genetics has been added. For the reader in possession of a first edition who wishes to know whether to splash out on a second, Figure P.1 may be helpful. (This is partly so that I can underline the fact that nothing in life, not even an author's preface, should be exempt from statistics!) This compares the two editions chapter by chapter in terms of their length in thousands of words.

The major division of the book occurs after Chapter 5, when material introducing statistics in drug development from historical, philosophical, technical and professional perspectives is succeeded by single-issue chapters. Figure P.2 shows even more clearly that much of the additional material has gone into chapters in Part 2. Apart from the wholly new chapter on pharmacogenetics in particular, chapters on baselines, measuring effects, intention to treat and missing data, equivalence, meta-analysis, sequential analysis, dose-finding, pharmaco-epidemiology and pharmaco-economics have had extensive additions.

It is appropriate for me to repeat a general warning about the book that Carl-Fredrik Burman has drawn to my attention. A number of the section headings contain statements of position. For example, Chapter 7 has a section, 'The propensity score is a superior alternative to adjusting for confounders than analysis of covariance'. You will get a very misleading impression of the message of the book if you take these as being *my* position. This is a book about issues and where such statements are made it is nearly always because, as is the case with this one, I wish to take issue with them.

I am grateful to Frank Bretz, Diane Elbourne, Paul Gallo, Oliver Keene, Dieter Hauschke, Nick Holford, Paul Johnson, Vincent Macaulay, Helmut Schütz, Helen Senn and John Sorkin for comments, and to Mateo Aboy, Nick Holford, Jerry Nedelman,

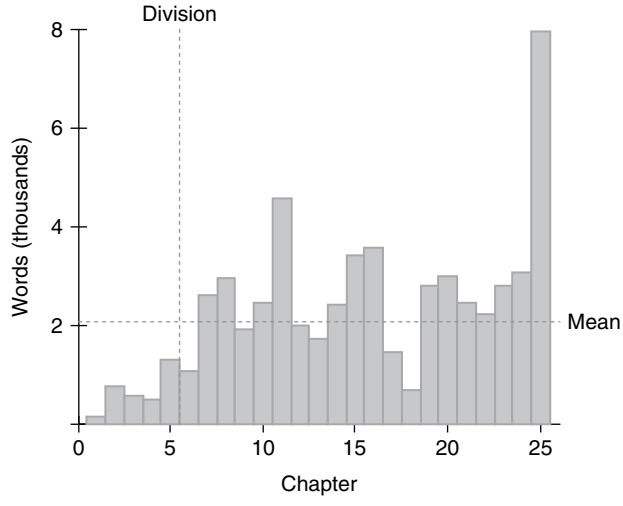

Figure P.2 Additional length (in thousand of words) of the second edition compared to the first. The mean increase (excluding chapter 25) is 2000 words.

Luis Pereira and Michael Talias for providing copies of their papers. Since the first edition, I have acquired several new co-authors whose work is reflected in this edition: my PhD students Dimitris Lambrou and Sally Lee and also, Pina D'Angelo, Frank Bretz, Carl-Fredrik Burman, Angelika Caputo, Farkad Ezzett, Erika Graf, Emmanuel Lesaffre, Frank Harrell, Hans van Houwelingen, William Mezanotte, Christopher Miller, Diane Potvin, Peter Regan and Nicoletta Rosati, whom I thank. I am also extremely grateful to Andy Grieve, for continued collaboration and to my PhD students Steven Julious, Andy Garrett for their work. I thank Kathryn Sharples, Susan Barclay, Beth Dufour and Simon Lightfoot at Wiley and also Len Cegielka and Cherline Daniel for work on preparing the book.

I continue to hope, of course, that this book will aid dialogue between statisticians and life-scientists within the pharmaceutical industry but also hope that it will contribute to a wider appreciation of the interesting challenges that statisticians within the pharmaceutical industry face and the seriousness with which they are met. I hope that the reader finds both stimulation and enjoyment in encountering these challenges.

References

Berger VW (2005) *Selection Bias and Covariate Imbalances in Randomized Clinical Trials*. John Wiley & Sons, Ltd, Chichester.

Brown H, Prescott R (2006) *Applied Mixed Models in Medicine*. John Wiley & Sons, Ltd, Chichester.

Chevret S (ed.) (2006) *Statistical Methods for Dose-Finding Experiments*. John Wiley & Sons, Ltd, Chichester.

Ellenberg S, Fleming T, DeMets D (2003) *Data Monitoring Committees in Clinical Trials: A Practical Perspective*. John Wiley & Sons, Ltd, Chichester.

Hauschke D, Steinijans V, Pigeot I (2007) *Bioequivalence Studies in Drug Development: Methods and Applications*. John Wiley & Sons, Ltd, Chichester.

Lui K-J (2004) *Statistical Estimation of Epidemiological Risk*. John Wiley & Sons, Inc., Hoboken.

Marubini E, Valsecchi MG (1995) *Analysing Survival Data from Clinical Trials and Observational Studies*. John Wiley & Sons, Ltd, Chichester.

Molenberghs G, Kenward MG (2007) *Missing Data in Clinical Studies*. John Wiley & Sons, Ltd, Chichester.

O'Hagan A, Buck CE, Daneshkah A, Eiser JR, Garthwaite PH, Jenkinson DJ, Oakley JE, Rakpw T (2006) *Uncertain Judgements*. John Wiley & Sons, Ltd, Chichester.

Parmigiani G (2002) *Modeling in Medical Decision Making: A Bayesian Approach*. John Wiley & Sons, Ltd, Chichester.

Senn SJ (2002) *Cross-over Trials in Clinical Research* (2nd edition). John Wiley & Sons, Ltd, Chichester.

Spiegelhalter DJ, Abrams KR, Myles JP (2003) *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. John Wiley & Sons, Ltd, Chichester.

Whitehead A (2002) *Meta-analysis of Controlled Clinical Trials*. John Wiley & Sons, Ltd, Chichester.

Whitehead J (1997) *The Design and Analysis of Sequential Trials*, revised 2nd edn. John Wiley & Sons, Ltd, Chichester.

Willan AR, Briggs AH (2006) *Statistical Analysis of Cost-effectiveness Data*. John Wiley & Sons, Ltd, Chichester.

Preface to the First Edition

The conscientious mathematician acts in this respect like the lady who is a conscientious shopper. Wishing to satisfy herself of the quality of a fabric, she wants to see it and to touch it. Intuitive insight and formal proof are two different ways of perceiving the truth, comparable to the perception of a material object through two different senses, sight and touch.

Polya, *How to Solve It*

When I first started work in the Swiss pharmaceutical industry in 1987, I had already been working as a statistician for twelve years: three years in the National Health Service in England and then nine years lecturing in Scotland. What struck me then was how much I still had to learn, not only about medicine, pharmacology and drug development in general, but also about my own subject, statistics. Working with other professionals, many of them experts in fields of which I knew nothing, called for a different approach to statistics in at least four ways. First, I had to examine what I 'knew' in order to establish what was useful and what was not. Second, I had to supplement it with knowledge of many fields of statistical theory which were new to me (design and analysis of cross-over trials was an example). Third, I had to pay attention to the scientific fields of my colleagues, in particular pharmacology, and try to work out the implications in statistical terms. Fourth, I had to repay my colleagues' willingness to explain their sciences to me by making my statistical points in terms that were clear to them.

In doing the latter, it became clear to me that there were many things which I myself had taken for granted which were debatable. I also came to the conclusion that many of the difficulties with statistical ideas lie deeper than statistics itself (an example would be notions of causality). Furthermore, I became convinced that statisticians pay their non-statistical colleagues a disservice if they try to gloss over genuine disagreements. This book is an attempt to present many of the statistical issues in drug development in a way which is comprehensible to life scientists working in drug development whilst avoiding false consensus. The emphasis will be on the intuition which Polya, although himself a distinguished mathematician, valued so highly. Although addressed to the life-scientist it is my hope that many statisticians, in particular those studying medical statistics or embarking on a career in drug development, will also find it useful. Above all I hope that it will help communication between the disciplines: a process by which the statistician stands to benefit as much as any other professional in drug development. I cannot pretend, however, to be objective on all issues and am not even sure of what such objectivity might consist. In my defence, however, I think that I may justly claim, that if all viewpoints are not given equal care and consideration, the reader will at least come away with a wider awareness that other views exist, than would have been the case in a more conventional approach.

The book is in two sections. The first, and by far the shorter section, is based on a statistics appreciation course which I gave a number of times to my colleagues at Ciba-Geigy. In addition to a brief introduction it consists of four chapters, each of which takes a different view, of statistics in drug development: historical, 'philosophical', technical and professional. I recommend that every reader will either read these chapters or satisfy her or himself that the material is familiar. The second and larger section is loosely based on a course (but less technical than that course) which I give to students on the Postgraduate Diploma in Statistics at Neuchatel University and consists of nineteen chapters of varying length which may be read in almost any order and consulted as considered desirable. Each chapter consists of a brief background statement and then a second section split into a number of issues. These are sometimes presented as true open issues and sometimes as positions which one might take on an issue. (Where the latter is the case, the reader should be warned that I usually *disagree* with the position taken.) At the end of each chapter are a number of references. Sometimes these will have been referred to explicitly but in some cases they are merely listed because they are useful additional reading.

When it comes to handling statistics, the life-scientist in drug development has two particular advantages over his or her colleagues elsewhere. First, (s)he will, through the nature of the work, come into frequent contact with statistical problems. Consequently, a basic familiarity with some essential concepts will be obtained. Second, technical statistical matters will be handled as a matter of course by the statisticians assigned to the various drug development projects. Life scientists working elsewhere will not always be so fortunate as to have resident experts whom they may consult. This provides a further justification for my decision to concentrate on issues rather than technicalities.

I find the subject matter fascinating, of course, but am aware that the reader will not always share my enthusiasm. I have tried to leaven the mix by adding chapter quotations and even the occasional joke or anecdote, ignoring Sterne's warning that: *Tis no extravagant arithmetic to say, that for every ten jokes, - thou has got an hundred enemies; and till thou hast gone on, and raised a swarm of wasps about thine ears, and art half stung to death by them, thou wilt never be convinced it is so.*

This book would have been impossible to write without the help of many statisticians and life scientists from whom I received my education at Ciba-Geigy (now merged with Sandoz to form Novartis). In particular, amongst statisticians, I have to thank my former colleagues Farkad Ezzet, Hans-Peter Graf, Andy Grieve, Walter Kremers, Gunther Mehring, Amy Racine and Jakob Schenker for many helpful discussions during my time at Ciba-Geigy, as well as the various members of my own group, Bernhard Bablok, Nathalie Ezzet, Friedhelm Hornig, Rolf Meinert, Erhard Quebe-Fehling, Peter Sacares, Denise Till, Elizabeth Wehrle and Albert Widmer for shared work over the years, and also others in Switzerland, the USA and elsewhere, too numerous to mention. I also learned a great deal from the life scientists with whom I collaborated and I would particularly like to thank Reto Brambilla, Giovanni Della Cioppa, Brigitte Franke, Francesco Patalano and Bill Richardson for sharing the joys and pains of drug development with me. I also owe a particular thank you to William Jenkins who, in introducing me to the problem of portfolio management, led me to a wider appreciation of the role of statistics in drug development, and to Anders Hove, Keith Widdowson, Marc Cohen, Bill Huebner, Ronald Steele of CIBA-Geigy and Peter Regan, formerly of Strategic Decisions Group for working on this problem with me. My UCL colleagues, Vern Farewell and Rebecca Hardy made helpful comments on various chapters as did Leon Aarons, Peter Bauer, Michael

Budde, Stephen Evans, Farkad Ezzet, Dieter Hauschke, Oliver Keene, Walter Kremers, John Lewis, Bill Richardson, Joachim Röhmel, Mark Sculpher and John Whitehead. I thank Lew Sheiner for permission to reproduce figure 22.1 from one of his papers. I am also grateful to Peter Bauer, Roger Berger, Michael Dewey, Farkad Ezzet, Nancy Geller, Andy Grieve, Miranda Mugford and Wendy Ungar for giving me access to (as yet) unpublished work of theirs and to Yadolah Dodge, professor of statistics at the University of Neuchatel and the students on the postgraduate diploma there for the opportunity to expound this subject. Thanks are also due to Guernsey McPearson for contributing quotations and other material.

This page intentionally left blank

Acknowledgements

With thanks to the following copyright holders for their kind permission to reproduce quotations from their publications:

Chapter 1
Brazzaville Beach, by William Boyd, published by Sinclair-Stevenson. Reprinted by permission of The Random House Group.

Chapter 6
The Lottery in Babylon, by Jorge Luis Borges, Translated by John M. Fein from Labyrinths, copyright © 1962, 1964, by New Directions Publishing Corp. Reprinted by permission of New Directions Publishing Corp.

Chapter 9
Roots Schmoots: Journey Among Jews, by Howard Jacobson. Reproduced with permission from Peters Fraser & Dunlop Group Ltd.

Chapter 22
Enderby Outside, by Anthony Burgess. © Estate of Anthony Burgess from Enderby Outside (1968), reproduced with permission.

Chapter 24
The Bear, Wodwo (1967), by Ted Hughes. Reproduced with permission from Faber & Faber.

This page intentionally left blank

1

Introduction

Ye maun understand I found my remarks on figures, whilk . . . is the only true demonstrable root of human knowledge.

Sir Walter Scott, *Rob Roy*

Statisticians know that words are important to statistics, yet surely their importance is not fully recognized.

William Kruskal

Opinions are made to be changed – or how is truth to be got at? We don't arrive at it by standing on one leg.

Lord Byron, *letter to Murray*

1.1 DRUG DEVELOPMENT

Drug development is the process not only of finding and producing therapeutically useful pharmaceuticals and of turning them into high-quality formulations of usable, effective and safe medicines, but also of delivering valuable, reliable and trustworthy information about appropriate doses and dosing intervals and about likely effects and side-effects of these treatments. Drug development is a process carried out by *sponsors* (mainly pharmaceutical companies) and its acceptability is ultimately judged by *regulators*. It is an extremely complex business and the risks are high, but the potential rewards are also considerable.

It takes many years for a project to reach development. First, basic research must be undertaken to validate concepts and mechanisms. Assessments of commercial potential for diseases and therapies are also needed and these will continue throughout the life of a project. Next, a lead compound must be identified for a particular indication. This will then be subjected to a battery of screening tests to assess its potential in terms of therapeutic activity. Back-up compounds will also be investigated. If a compound looks promising, it will also be evaluated from both safety and practical points of view. Will it be easy to formulate? How many steps are involved in the synthesis? How difficult will it be to manufacture in large-scale quantities? Before a treatment can go into development, not only must satisfactory answers have been obtained to all these questions but a viable pharmaceutical formulation permitting further study must be

Statistical Issues in Drug Development/2nd Edition Stephen Senn
© 2007 John Wiley & Sons, Ltd

available. This can be an extremely delicate matter, involving work to develop suitable solutions, pills, patches or aerosols as the case may be.

If and when a molecule is accepted into development, animal studies will be undertaken in order to check safety and to establish a dose at which studies in humans may be undertaken. Once basic toxicological work has been undertaken, 'phase I' may begin and the first such studies may start. These will be single-dose studies in which lower doses are tried first and cautiously increased until a maximum tolerated dose may be established. In many indications such studies are carried out on healthy volunteers, but where the treatment is highly aggressive (and hence intended for serious diseases) patients will be used instead. In the meantime, longer-scale toxicological studies with animals will have been completed. Pharmacokinetic studies in humans will be undertaken in which the concentration–time profile of the drug in blood will be measured at frequent intervals in order to establish the rate at which the drug is absorbed and eliminated. These studies together, if successful, will permit multiple-dose studies to be undertaken.

Once maximum tolerated doses have been established, phase II begins and dose finding studies in patients are started. This is usually an extremely difficult phase of development but, if the drug proves acceptable, the object is that preliminary indications of efficacy should be available and that a firm recommendation for doses and dose schedules should emerge. Once these studies have been completed, the pivotal phase III studies can begin. These have the object of proving efficacy to a sceptical regulator and also of obtaining information on the safety and tolerability of the treatment.

A successfully completed development programme results in a dossier – an enormous collection of clinical trial and other reports, as well as expert summaries covering not only the clinical studies as regards efficacy and safety but also preclinical studies and other technical reports as well as details of the manufacturing process. If successful, the package leads to registration, but even during the review process, phase IV studies may have been initiated in order to discover more about the effect of the treatment in specialist subpopulations, or perhaps with the object of providing data to cover price negotiations with *reimbursers*.

Regulatory dossier: A mountain of documents which takes a forest of trees to obscure the wood.

Once a drug has been launched on the market, the process of monitoring and 'pharmacovigilance' begins in earnest, since the drug will now be used by far more persons than was ever the case in the clinical trials in phases I to III, and rare side-effects, which could not be detected earlier, may now appear. Some further phase IV postmarketing studies may be initiated and further work extending indications or preparing new formulations may be undertaken.

1.2 THE ROLE OF STATISTICS IN DRUG DEVELOPMENT

There is no aspect of drug development in which statistics cannot intrude: from screening chemicals for activity to forecasting sales. Because the efficacy and safety of treatments

has to be judged against a background of considerable biological variability, all of the judgements of efficacy boil down in the end to a numerical summary of evidence whose message can only be understood with the help of the science of statistics. It is the norm that clinical trials are planned jointly by a statistician and a physician. Statisticians are also becoming more active in the shaping of projects as a whole. Furthermore, whereas in the past in Europe, the expert reports for a dossier would largely be a subjective qualitative assessment of evidence from individual trials, regulators increasingly expect to see quantitative summaries, so called meta-analyses. Hence statistics is increasing its empire throughout the drug development process.

For many years now, the Food and Drugs Administration (FDA) in the United States has employed statisticians to assist in its review process, and this is now gradually being imitated in Europe by various national regulatory agencies, although the European Medicines Agency (EMEA) has yet (as far as I am aware) to follow suit. Statistical issues have long been covered by FDA general guidelines and there are now separate specific statistical guidelines produced by the Committee for Proprietary Medicinal Products (CPMP) in Europe, and by the Japanese Ministry of Health and Welfare. The International Conference on Harmonisation (ICH) is also addressing the subject. There are also many national and international societies specifically for statisticians in the pharmaceutical industry. The exact duties of the statistician working in drug development are covered in some detail in Chapter 5.

1.3 THE OBJECT OF THIS BOOK

As implied by our chapter quotations, this book is about figures, words and opinions; more precisely, it is an attempt to give a largely verbal account of the various opinions that are held by those who figure. The purpose is not to present an authoritative prescription as to how to deal with each particular application of statistics in drug development. I myself am far too cynical to believe that such statements are possible. Instead, the object is to make the reader think about genuine statistical controversies in drug development. In many cases the issues are deep, and thinking about them will increase understanding about scientific problems of evidence and inference as they relate to pharmaceuticals and their role in medicine and health, whether or not the reader's thinking about them concludes in resolution of the issues in question to his or her satisfaction. I hope the process is not only beneficial but enjoyable.

In complete contrast to my first book, *Cross-over Trials in Clinical Research* (Senn, 1993, 2002), this book includes very few concrete examples. There are a number of reasons, of which two are particularly important. First, it is not my object to teach the reader how to *calculate* anything. There is no need, therefore, of data to calculate upon. Second, the book is deliberately minimalist and I have eliminated anything that interferes with proceeding to discussing the issues.

The scope of the book is statistics in drug development in humans. I have taken this to include issues concerning the value of drugs that we may develop (Chapter 24) and also in monitoring drugs that we have developed (Chapter 23). With the exception of these two chapters, however, the book as a whole is concerned with matters affecting clinical trials in phases I to IV. There are many other areas in which statistics in the pharmaceutical industry are important thus I have not covered, including, for example, chemometrics and classification of molecules in research, screening for

efficacy, toxicology, pharmaceutical development, stability testing, process optimization and quality control. If I have left these out it was partly because one has to drawn the line somewhere but mainly because what I know about them could be written on the back of an envelope. This does not mean I have left them out without regrets: I would have enjoyed learning about them, but it would have been unrealistic to have included them and, in any case, I wanted a book that reflected my actual experience of drug development. I hope that the book nonetheless serves to give an impression of how wide the subject is, but the reader should bear in mind that it is wider yet than covered in these pages.

1.4 THE AUTHOR'S KNOWLEDGE OF STATISTICS IN DRUG DEVELOPMENT

The fact that I have eliminated from this book the subjects about which I know nothing does not mean that I am equally knowledgeable about all those that remain. I think it is appropriate for me to warn the reader of strengths and weaknesses. If we exclude the chapters in Part 1, then chapters 6 through to 18 are strengths, as is chapter 22; chapters 19, 21 and 23 are weaknesses and chapters 20 and 24 fall somewhere between the two. Chapter 25 on pharmacogenetics, which is new, is a bit of an odd one out. My knowledge of genetics is poor. However, equally relevant to this chapter is an understanding of sources of variability and also of trial design, which are both areas in which I have researched extensively over the years. Extensive references, in many cases with recommendations for further reading, are given in each chapter.

1.5 THE READER AND HIS OR HER KNOWLEDGE OF STATISTICS

Ideally, nobody should study statistics who hasn't studied it already. The least stimulating aspect of the subject (I will not say the easiest since it has difficulties of its own), is the mechanics of calculation. Although many of the horrors of this topic have been eliminated by modern computing, it is inevitable that a first course in statistics will not avoid dealing with these algorithmic matters in some detail. (And those readers who find computing to be a horror in its own right have my sympathy.) Only in a second course, where many of the rudiments of 'how' have been answered, can the more interesting 'why', 'when' and 'whether to' questions be addressed.

> **Statistics**: A subject which most statisticians find difficult but in which nearly all physicians are expert.

Throughout this book it is assumed that the reader already has some basic familiarity with statistics, such as may be obtained, for example, by the excellent elementary text on medical statistics by Campbell, Machin and Walters (Campbell *et al.*, 2007). For a more detailed coverage I recommend Altman (1991) or van Belle *et al.* (2004), and

for a more advanced level Armitage and Berry (2001). Some basic familiarity with statistics in clinical trials as covered, for example, in the classic text by Pocock (1983) is also assumed. Other treatments of statistics in clinical trials that I can recommend are the broad but elementary book by Wang and Bakhai (2006), the mathematically more advanced text by Matthews (2006), or the more comprehensive treatment by Piantadosi (2005). As regards the drug development process itself, I have found Hutchinson (1993) extremely useful to give to my own students wanting a quick introduction, and useful guides to the major classes of drug are Youngson (1994) and Henry (1994). Nevertheless, there is no heavy reliance on this assumed knowledge and the reader's memory will be jogged from time to time regarding relevant matters.

1.6 HOW TO USE THE BOOK

Part 1 consists of four chapters that give a crash course on statistics in drug development by presenting four different perspectives of the matter: historical, methodological, technical and professional. The least authoritative of these is the first, Chapter 2, which gives the historical view and where I have had to rely on secondary sources via the expert commentaries of others (with the exception of the history of the *t*-test, (Senn and Richardson, 1994) and Fisher's involvement in trials in humans (Senn, 2006), where I have undertaken some original researches myself). Nevertheless, if one does not consider the history of a subject, however crudely, one is all too vulnerable to the myth of 'present perfect' and I felt that it was important to cover history. Consideration of past imperfect helps to introduce a sense of proportion.

> **Analysis plan**: A detailed description of the intended analysis written before un-blinding data in the pharmaceutical industry and somewhat later, if at all, elsewhere.

The crucial chapters are Chapters 3 and 4, which give the 'iron rations' of the subject, covering basic statistical notions of causality and experimentation and introducing the two major schools of statistics: the frequentist and the Bayesian. (However, the general issue of Bayesian versus frequentist methods is too big for me to tackle seriously, preoccupying as it does many of the finest minds in the statistical profession. It cannot be entirely ignored, however, and it does break surface at various points throughout the rest of the book. Historical and philosophical accounts are given in my book *Dicing with Death* (Senn, 2003).) Chapter 5 explains something of the duties and concerns of the medical statistician working in drug development and may help to set the scene for the issues that follow. Chapters 2–5 together make a suitable one-day introductory course to clinical trials. Indeed, as mentioned in the preface to the first edition, they grew out of a course I gave while still working at CIBA-Geigy, which I left in 1995, and I have frequently given such courses both before and since the appearance of the first edition in 1997.

The issues themselves are covered in Part 2 and are grouped in chapters by broad theme which may be read in almost any order (or not at all) as the reader wishes. The chapters from Chapter 19 onwards are rather more technical and specialized than

Chapters 6–18, which are closely related to a postgraduate course I gave for many years at the University of Neuchatel, and it may help to have read some of these earlier chapters first. The grouping reflects drug development more than it reflects statistics. For example, if a purely statistical arrangement had been envisaged, there would have been a chapter on random-effect models, with entries on meta-analysis, multicentre trials and n-of-1 trials, rather than separate chapters on these topics with sections dealing with random-effect models. Certain such statistical topics are therefore more profitably hunted down via the index rather than the table of contents. The topics chosen are also those that are *particularly* relevant to drug development. Thus, topics that are relevant only because they affect the whole of statistics, such as, 'what should the role of robust statistics be?', are scarcely touched on. A few of the chapters in Part 2 have technical appendices, where I felt that the matters covered were not, or were not easily, available in the literature. These may be safely ignored by all but the more statistically minded.

In some of the chapters the reader will come across terms highlighted in bold type. This is an indication that these are important concepts and is usually a sign that there is an entry in the glossary. An exception is Chapter 2, where I have also highlighted the names of important historical figures.

The chapters can be used in a number of ways: by junior statisticians in order to get a quick overview of issues affecting a particular type of trial they are working on for the first time, by physicians and other life-scientists working in the pharmaceutical industry to help them discuss statistical issues with their statistical colleagues, and by university departments as the basis for student seminars and journal clubs.

Finally, for those who find controversy unsettling, I can do no better than quote the heroine of William Boyd's *Brazzaville Beach*: 'I have taken new comfort and refuge in the doctrine that advises one not to seek tranquility in certainty, but in permanently suspended judgement.' From *Brazzaville Beach* by William Boyd, published by Sinclair-Stevenson. Reprinted by permission of The Random House Group Ltd.

References

Altman DG (1991) *Practical Statistics for Medical Research*. Chapman and Hall, London.

Armitage P, Berry G, Matthews JNS (2001) *Statistical Methods in Medical Research* (4th edition). Blackwell Science, New York.

Campbell M, Machin D, Walters S (2007) *Medical Statistics: A Textbook for the Health Sciences*, (4th edition). John Wiley & Sons, Ltd, Chichester.

Henry J (ed.) (1994) *British Medical Association: New Guide to Medicines and Drugs*. Dorling Kindersley, London.

Hutchinson DR (1993) *How Drugs are Developed*. Brookwood Medical Publications, Richmond on Thames.

Matthews JNS (2006) *An Introduction to Randomized Clinical Trials*. Arnold, London.

Piantadosi S (2005) *Clinical Trials: A Methodologic Perspective*. John Wiley & Sons, Inc., New York.

Pocock SJ (1983) *Clinical Trials, A Practical Approach*. John Wiley & Sons, Ltd, Chichester.

Senn SJ (1993) *Cross-over Trials in Clinical Research*. John Wiley & Sons, Ltd, Chichester.

Senn SJ (2002) *Cross-over Trials in Clinical Research* (2nd edition). John Wiley & Sons, Ltd, Chichester.

Senn SJ (2003) *Dicing with Death*. Cambridge University Press, Cambridge.

Senn SJ (2006) An early 'Atkins' Diet': RA Fisher analyses a medical 'experiment'. *Biometrical J* **48**: 193–204.

Senn SJ, Richardson W (1994) The first t-test. *Statistics in Medicine* **13**: 785–803.

van Belle G, Fisher LD, Heagerty PJ, Lumley TS (2004) *Biostatistics: A Methodology for the Health Sciences* (2nd edition). John Wiley & Sons, Inc., Hoboken.
Wang D, Bakhai A (eds) (2006) *Clinical Trials: A Practical Guide to Design Analysis and Reporting*. Remedica, London.
Youngson RM (1994) *Prescription Drugs*. Harper-Collins, Glasgow.

This page intentionally left blank

Part 1

Four Views of Statistics in Drug Development: Historical, Methodological, Technical and Professional

. . . when the usual symptoms of health or sickness disappoint our expectation; when medicines operate not with their wonted powers; when irregular events follow from any particular cause; the philosopher and physician are not surprised at the matter, nor are ever tempted to deny, in general, the necessity and uniformity of those principles by which the animal economy is conducted. They know that a human body is a mighty complicated matter, that many secret powers lurk in it, which are to us altogether beyond our comprehension . . .

David Hume, *An Enquiry Concerning Human Understanding*

This page intentionally left blank

2

A Brief and Superficial History of Statistics for Drug Developers

O that this all important science might become part of University Education
Florence Nightingale (writing about statistics)

2.1 INTRODUCTION

Statistics is the science of collecting, analysing and interpreting data. Statistical theory has its origin in three branches of human activity: first, the study of mathematics as applied to games of chance; second, the collection of data as part of the art of governing a country, managing a business or, indeed, carrying out any other human enterprise; and third, the study of errors in measurement, particularly in astronomy. At first, the connection between these very different fields was not evident but gradually it came to be appreciated that data, like dice, are also governed to a certain extent by chance (consider, for example, mortality statistics), that decisions have to be made in the face of uncertainty in the realms of politics and business no less than at the gaming tables, and that errors in measurement have a random component. The infant statistics learned to speak from its three parents (no wonder it is such an interesting child) so that, for example, the word *statistics* itself is connected to the word *state* (as in *country*) whereas the words *trial* and *odds* come from gambling and *error* (I mean the word!), has been adopted from astronomy

Because of its close relation to the mathematics of probability, it is worth, however, drawing one distinction between the science of *probability* as understood by statisticians and mathematicians and the science of *statistics* proper. The former starts with known or presumed laws and uses the calculus of probability to say something about the chance of given events occurring. (For example, given a fair die, what is the probability of a total score of exactly 8 in two rolls?) The latter starts with data and (given some knowledge of the circumstances under which they were obtained) tries to deduce something about the state of nature. (For example, given the results of a number of rolls of a die, can I say whether it is fair or not?) This latter problem is an example of what is known

Statistical Issues in Drug Development/2nd Edition Stephen Senn
© 2007 John Wiley & Sons, Ltd

as *inverse probability* (Dale, 1991) which, either directly or indirectly, is the concern of statistics.

Statistics is thus also the science of inferring facts about nature given the evidence available. Therefore, it is closely related to the branch of philosophy known as epistemology. It may surprise those who are unfamiliar with statistics to learn that statisticians and philosophers have worked on the same problems. For example, the nature of probability itself (is it a subjective feeling, a relative frequency or a propensity of systems?) is a matter for philosophical speculation.

In this chapter we shall follow the development of probability and statistics more or less chronologically from the 17th century as it appears in a study of the mathematicians who contributed to these developments. When coming to medical statistics, we shall then return in time to pick up some of the developments made by biologists and physicians.

2.2 EARLY PROBABILISTS

A number of 17th-century mathematicians, notably **Pierre Fermat** (1601–1665), **Blaise Pascal** (1623–1662) and **Christian van Huygens** (1629–1695), made important contributions to the theory of probability. Hacking (1975) has made the claim that probability emerged as a concept around about the middle of the seventeenth century, and then developed rapidly but was scarcely known in modern form before this time. The contribution of Fermat and Pascal through their correspondence of 1654 discussing problems put to Pascal by the French gambler the Chevalier de Méré, is particularly important and often taken as being the start of the subject. Much of these early researches into probability problems consisted of deriving mathematical laws for enumerating ways in which particular types of events may occur. One well known illustration of such a law, the binomial law, is given by Pascal's triangle (Pascal is the author of an elegant treatise on this triangle proving many results which, although already known, had not been proved before. See Edwards (2002) for a scholarly examination of Pascal's work.)

The first few rows of Pascal's triangle (which carries on for ever) are given in Figure 2.1. (Armitage, a famous medical statistician, gives an amusing account of a

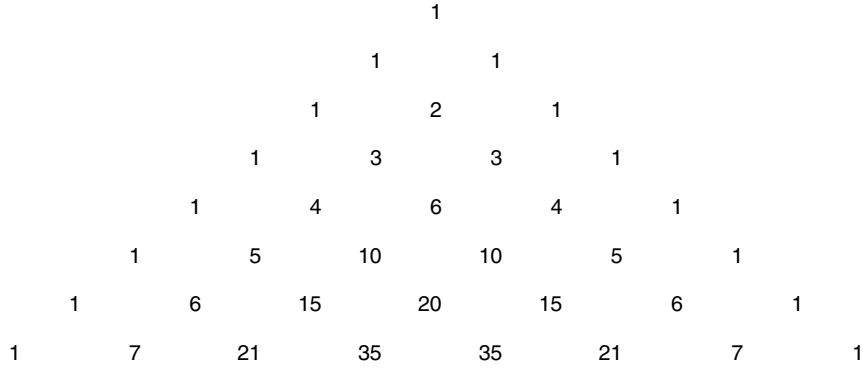

Figure 2.1 Pascal's triangle.

rather large Pascal triangle with 125 rows constructed by him, Godwin and Lindley, which was pasted to the wall of the office in which they worked during the Second World War. See Barnard and Plackett (1985).) Here, each number in a given row is the sum of the numbers nearest to it in the row above. For the first and last numbers in a row there is one such nearest number; for others there are two. For example, to use Pascal's triangle to enumerate the number of ways in which three heads may be obtained in five tosses of a coin we use the *fourth* entry in the *sixth* row (the first entry corresponds to 0 heads and the first row to 0 tosses) giving the answer 10. Since the total of all of the entries in the sixth row is 32 (corresponding to $2^{6-1} = 2^5$), then if on a single toss a head is just as likely as a tail, the probability of obtaining three heads in five tosses is 10/32. This general approach of calculating the ratio of favourable outcomes to total possible outcomes plays an important role in many probability calculations. Fermat's and Pascal's work can be regarded as a contribution to the theory of the binomial distribution, still applied today for studying binary outcomes. In the context of the clinical trial these might be dichotomies of the form survive/die, cured/not cured and satisfactory/unsatisfactory although, of course, unlike head/tail, such events would not be regarded as being equally likely. (The more general formula for the binomial distribution deals with such cases.)

2.3 JAMES BERNOULLI (1654–1705)

James (or Jakob, Jacques, etc) Bernoulli has been designated 'father of the quantification of uncertainty' (Stigler, 1986) and his *Ars Conjectandi*, published posthumously in 1713, is a landmark in the mathematics of probability. There are so many notable mathematicians among members of the well-known Bernoulli family of Basle (eight according to Bell (1953) but twelve according to Stephen Stigler and Boyer (1991)) that there is often confusion as to which is meant (uncertainty as to the father of the quantification of uncertainty?). Among statisticians, at least, *Bernoulli* without further qualification means, or ought to mean, *James* (the first James Bernoulli of note), who was Professor at the University of Basle from 1687 until his death in 1705. Figure 2.2, based on Bell (1953) and Boyer (1991) gives a pedigree. The Bernoullis with dashed boxes are those admitted to the pantheon by Boyer but not by Bell, although, rather confusingly, Bell charts Nicholas II but, perhaps showing a mathematician's prejudice against probabilists, does not count him among the great. However, not only is it thanks to Nicholas's actions in publishing his uncle James's work after his death, that *Ars Conjectandi* finally emerged to public scrutiny, but Nicholas himself made important contributions to the theory of probability as has now been made clear in Ander's Hald's masterly two-volume history of statistics (Hald, 1990, 1998). (Nicholas Senior's role in all this is merely to father a dynasty of mathematicians!)

James Bernoulli's skill in developing the methods of the calculus and his general familiarity with limits was put to a remarkable use in developing what we now call *Bernoulli's weak law of large numbers*. This sought to relate the observed relative frequency of successes in a large series of trials to the individual probability of success using a probabilistic law. It is an attempt to answer questions such as 'if the probability of obtaining a 5 in a single roll of a die is 1/6, in what sense may I expect that, given a large number of rolls, the proportion of 5s will be approximately 1/6?' Bernoulli's result is central to the science of statistics but its interpretation remains controversial.

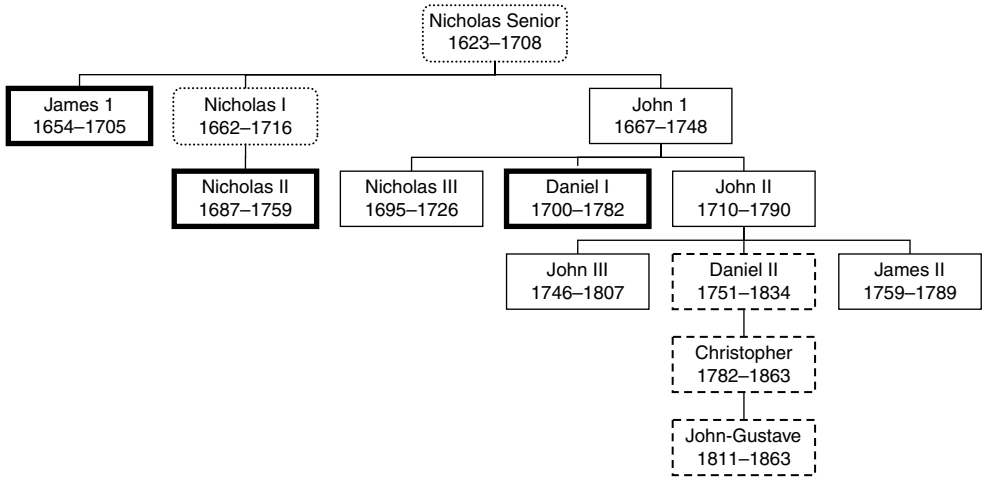

Figure 2.2 Family tree of the Bernoullis (based on Boyer, 1991). Nicholas Senior and Nicholas I were not mathematicians. James I, Nicholas II and Daniel I are particularly important in the history of statistics. Daniel II, Christopher and John-Gustave are not included by Bell (1953) but are included by Boyer (1991) and Stigler (1986).

2.4 JOHN ARBUTHNOTT (1667–1753)

A famous example of applied statistics in the 18th century is Arbuthnott's paper of 1710 *An argument for Divine Providence, taken from the constant regularity observ'd in the births of both sexes* (Arbuthnot, 1710), which used the fact that the probability was minute that the yearly observed excess of male compared female births (if christenings could be accepted as a measure) in London from 1629 to 1710 was due to chance (if it were assumed that male and female births were equally likely) as an argument that the Almighty must be intervening for good in the affairs of mankind by providing a greater number of the frailer sex. (Here frailer must be understand in its statistical sense! Women are the stronger sex.) Arbuthnott's analysis was improved upon by a **Bernoulli** (**Nicholas II**, 1687–1759, this time, nephew of James, and mentioned above) who, however, may have held rather different opinions regarding the value of this demonstration of Divine Providence.

John Arbuthnott earned his living as a teacher of mathematics and translated Huygens' (see above) treatise on probability into English. Unlike most mathematicians and statisticians, he evidently had a keen appreciation of practical economics as he then studied medicine and became a fashionable physician.

2.5 THE MATHEMATICS OF PROBABILITY IN THE LATE 17TH, THE 18TH AND EARLY 19TH CENTURIES

These developments are too numerous to cover in detail, but a few lines of work in particular are extremely important. The first is the development of the mathematics of insurance, associated in particular with the rise of commerce in Holland and in Britain. The Dutch politician **John** (Jan, Johan, etc) **de Witt** (1625–1672) and the English

astronomer **Edmund Halley** (1656–1742) introduced methods for the calculation of annuities, the latter constructing a life table from records of vital statistics of the city of Breslau (Wroclaw). Halley's work was extended by the French-born English mathematician **Abraham de Moivre** (1667–1754), who also did important work on the binomial distribution, and by **John Simpson** (1710–1761), who was a weaver who supplemented his income by teaching mathematics. (Simpson is also honoured eponymously among mathematicians through *Simpson's rule*, a method for approximating the area under a curve. In drug development, areas under the curve are commonly calculated in bioequivalence and dose-proportionality studies, although the simpler trapezoidal rule is more usually employed.)

Modern survival analysis, an important tool of statistical analysis in drug development today, can be regarded as having its origins in the work of De Witt, Halley, de Moivre and Simpson and other mathematicians of the 17th and 18th centuries who worked on life tables and annuities.

A further important impetus for the development of probabilistic study came from the study of 'errors' in observation in astronomy and surveying, where scientists began to realize that, although individual errors were not predictable, their distribution *en masse* followed particular 'laws'. The study of this phenomenon and methods for dealing with it led to much important work by mathematicians, such as on the Normal distribution and on methods for combining observations, in particular the method of least squares. Among the many mathematicians who worked on these problems, particularly notable are the Swiss mathematician **Leonhard Euler** (1707–1783), the German astronomer and cartographer **Tobias Mayer** (1723–1762), and especially the French mathematicians **Pierre-Simon Laplace** (1749–1827) and **Adrien-Marie Legendre** (1752–1833) and the German mathematician **Carl Friedrich Gauss** (1777–1855). The Normal distribution has a central place in the subject of statistics today and the method of least squares is one of the most common methods of statistical estimation.

2.6 THOMAS BAYES (1701–1761)

A woman aged 40–44 at parturition is more likely to have a baby suffering from Down's syndrome than one aged 20–24, but to conclude, *therefore*, that any newborn baby suffering from Down's syndrome is more likely to have been born to a woman aged 40–44 than to one aged 20–24 is quite false. To solve this sort of problem we use one of the most important theorems in statistics namely *Bayes' theorem*.

Suppose we label the two classes of mother as 'young' and 'old' (ignoring all other age groups) and suppose that the probability of an old mother having a child with Down's syndrome is 1 in 40 ($=0.025$) and for a young mother it is 1 in 200 ($=0.005$). Suppose, furthermore, that the probability of any baby at all (irrespective of chromosomal status) having a young mother is 3 in 10 (0.3) and of having an old mother is 1 in 100 (0.01). Use of Bayes' theorem then shows that the relative probability of a Down's syndrome baby having a young mother is $(0.005 \times 0.3)/(0.025 \times 0.01) = 6$. Thus the Down's syndrome baby is six times as likely to have had a young mother as an old mother. (Of course the probabilities used in calculating this example are purely illustrative and should not be taken seriously.)

Until relatively recently, little was known about the Bayes of *Bayes' theorem*, but it is now known that he was a nonconformist minister in Tunbridge Wells who was

elected a Fellow of the Royal Society in 1742. His *An essay toward solving a Problem in the Doctrine of Chances* (1764) was published after Bayes' death by his friend Richard Price in 1763 (see Barnard, 1958). There is a very vigorous modern school of statistics which is referred to as Bayesian because of the importance of Bayes' theorem in its approaches.

> **History**, n. An account mostly false, of events mostly unimportant, which are brought about by rulers mostly knaves, and soldiers mostly fools.
>
> Ambrose Bierce.

2.7 ADOLPHE QUETELET (1796–1874)

The Belgian Adolphe Quetelet is an important, if curious, figure in the history of statistics: mathematician, astronomer, man of letters and, ultimately, statistician. He was the first to attempt a systematic application of statistical methodology to what we should now call sociology and although, in Stigler's assessment, 'Quetelet did not accomplish a great deal towards this end – in some respects he failed totally' (Stigler, 1986, p. 161) he had an extremely profound effect on the generations after him. Our chapter quotation is taken from marginal notes by Florence Nightingale on a book of Quetelet's (Diamond and Stone, 1981) (she read and re-read Quetelet); he was an important influence on Galton (see below) and the claim has even been made that James Clerk Maxwell gained the confidence to develop statistical mechanics to explain the behaviour of gases from the example of Quetelet (Gigerenzer *et al.*, 1989), inspired by the ability in sociology to show order arising out of chaos through the aggregation of the random behaviour of numerous individuals.

2.8 FRANCIS GALTON (1822–1911)

Galton was an eccentric amateur scientist who studied both medicine and mathematics and made important contributions to a number of fields (Forrest, 1974). Among his many discoveries and inventions were a method for classifying fingerprints, a machine for extracting wave energy, barometric maps (he was the discoverer and namer of the meteorological phenomenon of the anticyclone) and a feminine beauty map of Britain (whenever he passed a woman in the street he would make a mark on a paper cross he carried in his pocket recording her attractiveness or otherwise). He also made extensive studies in genetic heritability and was able to discover empirically the phenomenon of *regression* (now easily demonstrated mathematically).

Galton discovered that tall parents tend to have children who, while taller than average, are nevertheless smaller than their parents. They thus *revert to mediocrity* or *regress to the mean*. This phenomenon of regression is widespread and general so that, for example, patients who have been selected for a clinical trial on grounds of high blood pressure may, *on purely statistical grounds*, be expected, other things being equal, to show an improvement in blood pressure when measured again. This omnipresent, but widely ignored, feature is an important reason for the unreliability of uncontrolled